

Author comments in response to Reviewers' comments on “The Met Office Unified Model Global Atmosphere 8.0 and JULES Global Land 9.0 configurations” by Willett et al.

The Authors would like to thank both the reviewers for taking the time to review the manuscript and providing valuable comments. We believe that their contribution has resulted in a much improved manuscript.

The reviewers' comments are copied below and our response to them included after each individual comment in red

Reviewer 1.

The paper is a documentation of the latest model upgrade at the MetOffice. It is exceptionally well written and could almost be published as is. Since this is the documentation of a specific model, it is almost impossible to review it as an external, as a lot of the information is mainly relevant for the specific model. However, everything sounds sound and reasonable and I only have minor suggestions:

RESPONSE: Thank you. We have endeavoured to improve the paper in response to your comments and the comments of the other reviewer.

L43: “details the all”

RESPONSE: fixed

L69: Maybe add a reference on ENDGame?

RESPONSE: The main ENDGame reference, Wood et al, 2014 is now referenced later in the paragraph.

Section 2 reads a lot like a textbook, and it is not clear how much of this is actually needed. It does not serve as a full model documentation. Yet, it includes input that is not relevant for the model update. What is your aim of the section?

RESPONSE: The intention of section 2 is to give a referenceable description of the model without having to iteratively refer back to the previous papers in the series. This approach has been agreed with editors and has been the practice for all the previous papers in the series (Walters et al, 2019, 2017, 2014). To make this approach explicit the following have been added to section 1. "The primary aims of this paper are to provide a standalone scientific description of GA8GL9 (and hence a single reference) and to describe how GA8GL9 differs from the previous configuration, GA7GL7."

L723: “This its”

RESPONSE: fixed

L882: “shown shown”

RESPONSE: fixed

L896: “it primarily”

RESPONSE: fixed

Figure 7: I do not understand the figure and I get the feeling that it would be better to just include the information in the text.

RESPONSE: The figure shows the bias in temperature at 100hPa and humidity at 70hPa relative to ERAI and MERRA. The green rectangle, which was not explained in the caption, represents the "desirable" range of values for the temperature and humidity errors, and was included by default using our standard assessment tools. We have removed it from the figure to aid clarity and because it is not relevant to the point that the figure is meant to make.

L910: "and a changes"

RESPONSE: fixed

A message to the editor regarding L916: It is impossible to review the code with such a paper. It is therefore pointless to collect the code.

Reviewer 2.

Review of "The Met Office Unified Model Global Atmosphere 8.0 and JULES Global Land 9.0 configurations" by Willett et al.

This is a model description paper of a new version of the UK Met Office Unified Model (GA8GL9). This is a world-leading model for NWP and climate applications, and therefore this is an important paper. The paper contains a comprehensive description of the updates relative to the previous version of the model. However, I think a number of changes are needed to the manuscript before finalisation.

A large part of Section 2, which contains a description of the model, is copied directly from Walters et al. (2019 <https://doi.org/10.5194/gmd-12-1909-2019>), including many sentences which are copied verbatim, and some others which are copied with only light edits. All of the subsection titles in Section 2 are the same in both papers. The authors should instead either omit the copied text and just cite Walters et al., or summarise it in their own words if needed, focussing on general description of changes relative to GA7GL7.

RESPONSE: The intention of section 2 is to give a referenceable description of the model without having to iteratively refer back to the previous papers in the series. Furthermore, where the model is unchanged from previous release, we have largely left the text of the description unchanged. This approach has been agreed with editors and has been the practice for all the previous papers in the series (Walters et al, 2019, 2017, 2014). To make this approach explicit the following have been added to section 1.

- "The primary aims of this paper are to provide a standalone scientific description of GA8GL9 (and hence a single reference) and to describe how GA8GL9 differs from the previous configuration, GA7GL7."
- The footnote "Where the configuration remains unchanged from GA7GL7 and its predecessors, Sect. 2 contains material which is unaltered from the documentation papers for those releases (i.e. Walters et al, 2019, 2017, 2014, 2011)".

Section 3 is long and comprehensive and contains a description of all the changes made in the latest version of the model. I have the impression that the authors have collected a set of changes from a version control system, and described them all here. I got the impression that the section titles used in Section 3 were taken directly from a repository or issue tracker without any editing. I suggest editing the titles so that they are clearer to an outside audience. While the description of some changes is quite clear, the description of others is rather difficult to follow, and I suspect might only fully make sense to readers who are familiar with the code. Moreover while all the sections describe the changes, some of the sections lack a full scientific justification. I encourage the authors to shorten discussion of technical details, make the text more accessible to outside readers, and add more justification and results of analysis to motivate the changes. Walters et al. (2019) include more such justification in the corresponding section of their paper. Parts of this section of the manuscript also contain many grammar errors which could be picked up by proofreading (I flag a few in my comments but there are many others).

RESPONSE:

* Where necessary, the titles of the model changes have been modified to remove reference to the names of model parameter or model options and make them more physically descriptive.

* The text has been checked to ensure that the descriptions are clear and, where necessary, modifications have been made. Any specific comments about section 3 have been addressed.

* For many (14) of the science changes in Sect. 3, the text describing results and/or motivation has been expanded to varying degrees. Four figures showing results have also been added to Sect. 3.

* The text has been checked for grammatical and spelling errors, and corrected where necessary.

While Section 3 of the manuscript is very long, it only contains a single paragraph on tuning – I suggest devoting more attention to this.

RESPONSE: The tuning section has been very substantially expanded as discussed in response to a later comment.

Also the evaluation section of this manuscript is also very short – I suggest that the authors consider including more evaluation results in the manuscript.

RESPONSE: Most of the assessment of GA8GL9 was done against the branch configurations GA7.1GL7.1 for climate and GA7.2GL8.1 for NWP - this is discussed in section 4. This evaluation is presented in Xavier et al, 2024 <https://doi.org/10.62998/uzui3766> which is referenced in Section 4. We do not wish to repeat this evaluation here as it does not compare GA8GL8 with the previous trunk configuration, GA7GL7, and because having three different controls (GA7GL7, GA7.1GL7.1, GA7.2GL8.1) would be messy and confusing. We have, however, added an addition figure and some additional text to this section.

Is it really the case that the UM source code cannot be publicly shared? Doing so would seem to align with the UKRI policy on making the results of publicly-funded research publicly accessible (<https://www.ukri.org/what-we-do/supporting-healthy-research-and-innovation-culture/open-research/>). The source code for several other climate models internationally is now publicly

available, for example: <https://github.com/ESCOMP/CESM>, <https://github.com/E3SM-Project/E3SM>, <https://github.com/EC-Earth>, <https://gitlab.com/cccma/canesm> among others. Parts of the manuscript feel as though they are written for readers who have access to the code – including references to particular parts of the codebase, and description of bug fixes – so it really would make more sense if readers could access the associated code.

RESPONSE: The licensing of the UM and JULES source code is outside of the control of the authors and cannot currently be publicly shared. It is available for research and operation use under licence agreement.

The authors should check that all the acronyms they use are defined on first use.

RESPONSE: The acronyms should all now be defined on first use.

It would be helpful to have more context on the planned uses of this model in the intro. Will this model be the basis of the UKMO CMIP7 climate model? Will it be used in NWP operations?

RESPONSE: The text describing the applications of GA8GL9 and GC4 has been expanded. It now reads "GC4 has been used for operational global NWP at the Met Office since May 2022, but there are no plans to use it for operational climate applications. GA8GL9/GC4 is used as a baseline for future model developments (e.g. Lock et al. (2024)) and for other research activities. The next GA and GL configuration, GAL9, will build upon GA8GL9."

Specific comments:

Abstract: Were any of these versions used in UKESM (CMIP6 or CMIP7 versions?)

RESPONSE: The applications of GA7.1 and GA7.2 are now listed in parantheses.

Ln 31: Can you name the CMIP6 models? UKESM1-0-LL and HadGEM3-GC3.1?

RESPONSE: A list of climate configurations that use GA7.1 has been added.

Ln 35: If GL8 is a global land configuration why does it include changes in sea-ice drag?

RESPONSE: "Global Land" is admittedly a misnomer in that it includes not just land surface processes but also the surface exchange over sea and sea-ice. A footnote has been added where "Global Land" is first mentioned to explain this. Furthermore, the title of Sect. 2.11 has been changed to "2.11 Surface flux exchange, land surface and hydrology: Global Land 9.0" and the description of GL9 in this section has been expanded to describe how the surface exchange over sea and sea-ice are represented.

Ln 36: Can you explain what 'aggregate tile version' means?

RESPONSE: We have added a brief description of the aggregate tile and referenced Walters et al, 2017 where this is approach is discussed.

Ln 56: What does 'held at the north and south poles' mean?

RESPONSE: This has been slightly reworded and is hopefully clearer

Ln 116-123: It would be helpful to refer forward to Section 2.10 here. On reading this section I didn't know how the aerosol size distribution was modelled, even though this is described later, in Section 2.10.

RESPONSE: Agreed. A reference to section 2.10 has been added.

Ln 134-135: Consider also citing the original articles describing the McICA scheme, such as Pincus et. (2003), <https://doi.org/10.1029/2002JD003322>

RESPONSE: A reference to Pincus et al, 2003 has been added.

Ln 232-234: The meaning of the sentence “There are additional non-local fluxes of heat and momentum in order to generate more vertically uniform potential temperature and wind profiles in convective boundary layers” isn’t clear to me. Does this mean that the model includes these non-local fluxes in order to achieve more vertically uniform potential temperature and wind profiles in convective boundary layers? If so, why?

RESPONSE: This has been rewritten to give a clearer description. It now reads "In convective BLs, the scheme includes additional non-local fluxes of heat and momentum that represent the efficient mixing by convective thermals. These generate the vertically uniform potential temperature and wind profiles seen in observations."

Ln 318-323: It wasn’t clear to me why lakes are represented as a canopy. Doesn’t this imply that the lake interacts with the underlying surface in the model, like a tree canopy? Is the same approach used for large lakes like the Great Lakes or Caspian Sea? Does the approach change in conditions where lake ice would be present?

RESPONSE: The text has been adjusted to make it clear that we are following the approach described in the original JULES paper (Best et al, 2011 <https://doi.org/10.5194/gmd-4-677-2011>). A statement that the lake canopy is essentially decoupled from the underlying surface has been added. Strictly speaking the definition of the land-sea mask, and hence the treatment of the larger lakes, is application specific and outside of the definition of GA8GL9; however, a footnote has been added that states that the larger lakes are treated as sea points in all applications.

Ln 345-348: If freshwater outflow from inland basins, distributed evenly across all sea outflow points, is an important component of the thermohaline circulation, isn’t this a concern? Physically in the real world there is no such transfer of freshwater – doesn’t freshwater only leave inland basins by evapotranspiration?

RESPONSE: This process only happens when the soils becomes saturated and cannot hold anymore water. The runoff from saturated soils would normally be put into the rivers, but for inland basins this would just return the water to the same point. We agree that the existing description is too brief and therefore we have changed it to "If the soil at inland basin points becomes saturated and hence unable to hold more water, the resultant runoff is pragmatically redirected evenly across all sea outflow points. The resulting increase in the river outflow is always very small and does not affect the ocean salinity structure in any significant way, but importantly this process ensures global water conservation which is an important criterion for global climate models."

Ln 344-353: How is runoff over ice sheets handled?

Liquid water (from rain or snow melt) on land ice is passed the surface runoff and hence to the rivers in a similar manner to other surface types, although this term is usually very small. Run off from land ice is now described as part of a newly added description of land ice in Sect. 2.11.

The interaction of run off with the rivers is also discussed in Sect. 2.11. and a comment to state this includes land ice has been added.

Ln 368-369: The preceding text gives a physical justification for the use of stochastic physics schemes, but this seems to be contradicted here by the statement that their formulation ‘lacks a sound physical basis’.

RESPONSE: Agreed that this could be read as contradictory. The text now reads: "Despite the positive impact of model error stochastic physics schemes on EPS and climate model performance by making probability forecasts more statistically reliable and reducing the error of the ensemble mean, their formulations fundamentally add noise which degrades forecast skill when run in deterministic mode (Sanchez et al., 2016). For these reasons, these schemes are not used in deterministic forecast systems, which are designed to forecast the best possible single prediction of the atmosphere's future state."

Ln 389-391: Are these GMED tickets publicly accessible somewhere? If not, perhaps it is not so helpful to list them here. If they are available, give a reference or URL.

RESPONSE: The GMED tickets are not publicly accessible, but they are accessible to many users of the UM and JULES. The GMED ticket numbers are arguably useful for those with access to this system and, in the opinion of the authors, do not detract from the readability of the paper for the wider audience. On balance we feel that their inclusion is beneficial. This approach is consistent with Walters et al, 2019. The following has been added to section 1 to explain their inclusion - "The development of these changes is documented using “trac” issue tracking software; for consistency with that documentation, we have included the trac ticket numbers with each change." References to the GMED ticket numbers within the text have been removed apart from in the title of each change, the list of changes that were included in GA7.1 and GA7.2 given at the start of section 3, and the author contribution section.

Ln 515-534: This paragraph doesn't say explicitly what the new value of puns is. I re-read the paragraph and gathered that it is 1.0, but this wasn't clear on first reading. Also the phrasing 'value for puns of 0.5 has been used in GA7' makes it sound like GA7 is the topic of the current paper. I suggest '0.5 was used in GA7'.

RESPONSE: Changed to "value for puns of 0.5 was used prior to GA8". The following has also been added to ensure it is clear what value is used in GA8 - " For these reasons the value of puns is increased to 1.0 in GA8."

Ln 539: 'the time-damping is increments are defined as' – grammar error – delete 'is'.

RESPONSE: fixed.

Figure 1: x-axis should be labelled 'Fraction of time'.

RESPONSE: Label changed to "Fraction of time"

Ln 535-536: This text explains that the convective scheme is intermittent, but it doesn't explain why this is unrealistic. Isn't convection intermittent in the real atmosphere? Are the authors arguing that the convection was unrealistically intermittent compared to the real atmosphere?

RESPONSE: We consider that the intermittency is essentially a numerical artefact and is undesirable and unphysical. The text has been changed to make this clear, i.e. "The convection scheme can sometimes become intermittent in time when the effects of convection on the

environment on one timestep can spuriously prevent it being triggered on the subsequent timestep; this is ultimately caused by the time-explicit implementation of the convection diagnosis and closure. Although real convection can be highly variable in time, the intermittency is considered to be unphysical because when averaged over the spatial scales of a model gridbox, real convection does not typically appear and disappear on the timescale of a model timestep."

Ln 558-559: Check grammar "and showed that not only diffusive but also that it was only computational stable"

RESPONSE: Changed to "not only is it diffusive ..."

Ln 562: occasion -> occasional

RESPONSE: fixed

Ln 581: dimension -> dimensional

RESPONSE: fixed

Caption to Table 6: 'season mean diurnal cycles' -> 'seasonal mean diurnal cycles'

RESPONSE: fixed

Ln 625: single level -> a single level

RESPONSE: fixed

Ln 647: Not clear what 'otherwise it will be ignored' is referring to.

RESPONSE: The point about the convection scheme not supporting dry convection is not essential for the description of this change and has been removed.

Ln 680: The results -> These results

RESPONSE: fixed

Ln 685: 'CMT' is not defined.

RESPONSE: CMT is now defined in section 2.9.

Ln 690: What is 'ProgEnt'?

RESPONSE: this abbreviation for the prognostic based entrainment was introduced in section 3.5.3. To make this section clearer "ProgEnt" has been changed to "prognostics based entrainment (Sect. 3.5.3)"

Ln 694-702: Some comment on the importance of freezing convective rain would be helpful. Does this change reduce model biases? Was it done to better simulate freezing rain at the surface?

RESPONSE: The main motivations were to avoid convective rain freezing over a single level, which constitutes a vertical resolution dependence, and for consistency with the treatment of falling snow. The text has been expanded to clarify these points and also note that the impact is small.

Ln 715-720: Can you provide any more justification for the change in refractive index of BC?
Note that there are other estimates which have even higher absorption, such as Bescond et al. (2016): <https://doi.org/10.1016/j.jaerosci.2016.08.001>

RESPONSE: This change was added in GA7.1 and hence included in HadGEM3-GC3.1 and UKESM1. As stated in the introduction, GA8 reintegrates the changes made at GA7.1 back onto the trunk. This change in the context of GA7.1 and UKESM1 is discussed in Mulcahy et al, 2018 <https://doi.org/10.1029/2018MS001464>. The text has modified to state explicitly that we are following Mulcahy et al, 2018. As further justification the following has also been added "Mulcahy et al. (2018) demonstrated that in GA7.1 this update to BC properties greatly reduced the bias in absorption aerosol optical depth relative to AERONET (Aerosol Robotic Network) observations." Noted that there are a large range of values for the refractive index of BC, but the reasons discussed we have used we have used the middle value from Bond and Bergstrom (2006) which was a comprehensive study that provided a well-respected benchmark.

Ln 744-74: Could the authors add a few more words to explain why a scaling on PMOA was applied by Gantt et al. (2012), but is not applied here?

RESPONSE: Brief reasoning for this has now been added. Now reads "Gantt et al. (2012) applied a global scaling factor of 6 to the diagnosed emissions above. However, following Mulcahy et al. (2020), who found that the global emissions of PMOA in UKESM1 compared well with Gantt et al. (2012) without the scaling, we do not apply the scaling here."

Ln 748-758: Is volcanic aerosol diagnosed or prescribed? Section 2.10 indicates that stratospheric aerosol is prescribed, but this section suggests that it is simulated based on emissions. Or is this for stratospheric aerosols in the troposphere? Or are stratospheric volcanic aerosols simulated in the model but not used in the radiation scheme? Clarify.

RESPONSE: This has now been clarified. It now reads "The volcanic SO₂ emission input data that represents the continual degassing from volcanoes into the troposphere (rather than explosive injection into the stratosphere as discussed in Sect. 2.10) is updated ..."

Ln 757-759: This section should reference Section 2.2 where the loop structure is explained.

RESPONSE: Agreed. Section now starts "As described in Sect 2.2, ..."

Ln 782: doubled counted -> double counted

RESPONSE: fixed

Ln 810: convective parcel -> the convective parcel

RESPONSE: fixed

Ln 829-839: The description of the tuning process for the model is very brief. It is helpful to have a list of the parameters that were tuned and some discussion of the target metrics, but it would be good to have a more detailed discussion, given the importance of the tuning process. What were the target metrics? Was any kind of systematic tuning process applied? There is now an extensive literature on model tuning approaches – for example: <https://doi.org/10.1175/BAMS-D-15-00135.1>, DOI: 10.1126/sciadv.adf2758, <https://doi.org/10.5194/gmd-10-1789-2017>. Perhaps there were computational or people time limitations which prevented a more systematic approach to tuning, but it would be good to briefly discuss such approaches, and perhaps speculate on whether they might be applied to the UM in the future. Note that Walters

et al. (2019) include a much more comprehensive description of the tuning of the previous version of this model. One other related question - are any of the model parameters scale-independent?

RESPONSE: Agreed. The tuning section has been substantially expanded. It now includes a discussion of the model acceptance criteria, testing strategy and an explanation of tuning parameter selection. The reason for each parameter change is discussed and a figure showing the effect of the tuning on tropical temperatures has been added. There are only a couple of parameters that depend on resolution - a comment on this has been added.

Ln 852: cooling relative -> relative cooling

RESPONSE: fixed

Figure 3, caption: Why use ERA-Interim for model evaluation rather than ERA5? ERA5 generally has lower biases.

RESPONSE: Although it would undoubtedly be preferable to evaluate against ERA5 rather than ERA-Interim, at the time of the assessment of GA8GL9 and indeed at the time of writing, ERA5 had not been included in our standard assessment tools. This update is planned but has not yet been done. Figure 3 does, however, include some estimate of observational uncertainty based on a range of data sources.

Ln 890: his counteracts -> this counteracts

RESPONSE: fixed

Figure 7: What does the green rectangle represent?

RESPONSE: The green rectangle represents the "desirable" range of values for the temperature and humidity errors, and was included by default in our standard assessment tools. We have removed it from the figure to aid clarity and because it is not relevant to the point that the figure is meant to make.

Figure 8: are proportion to the change in RMSE -> are proportional to the change in RMSE

RESPONSE: fixed

Ln 907: computation stability -> computational stability

RESPONSE: fixed

Ln 978: lower Limit the CAPE timescale -> lower limit to the CAPE timescale

RESPONSE: fixed